Visual Validation of Estimated Parametric Models of Power Systems Under Ambient Conditions*

Urmila Agrawal
University of Wyoming
uagrawal@uwyo.edu

John W. Pierre
University of Wyoming
pierre@uwyo.edu

Abstract

With the advent of synchrophasors, a number of measurement–based algorithms have been developed to estimate system modes using ambient data. The accuracy of these mode estimates depends on a number of factors, such as selection of the system order and data channels. To validate these mode estimates, two methods are proposed in this paper based on how well they agree with the observed data using the estimated parametric system model and observed data. In the proposed methods, the contribution of each mode estimate is analyzed to obtain a combination of system mode estimates to represent the model. The methodology of the proposed methods are illustrated using least–squares ARMA (LS–ARMA) method, which gives a parametric system model estimates. Results obtained by implementing the proposed methods on measured and simulated data validate the effectiveness of the proposed methods.

1. Introduction

Small–signal stability is of great significance for maintaining reliable power systems operations. A useful indicator of the margin of this stability is the damping ratio of system modes [1]. With the advent of phasor measurement units (PMU), a number of measurement–based algorithms have been developed to estimate system modes using ambient data that can continuously monitor damping ratio of system modes and provide critical information on the system stability. Some of these mode–estimation algorithms are described in [2–8]. The accuracy of these estimates depend on a number of factors such as selection of the input variables required by the algorithm, for example model order selection for least squares auto–regressive moving average (LS–ARMA) method, the data quality and observability of mode of interest in the selected data channel, etc. Because of this, it becomes necessary to analyze how well these mode estimates describe the system, which can be done based on how well they agree with observed data and thus can help with the proper selection of data channels and input variables required by mode estimation algorithms.

Some of the work describing the performance of mode estimators are presented in [5, 9–12]. Papers [9, 10] are based on the residual whiteness testing of the prediction errors, [5] provides error bounds on the estimates, and [11, 12] provides a confidence interval of the estimates. In this paper, a different approach is followed, based on which two methods are proposed to validate these mode estimates by analyzing how well an estimated parametric system model matches with the observed data. These methods can be applied to any mode–estimation algorithm that can provide parametric estimates of a system model, for example N4SID subspace methods that give estimates of the state space representation of a system [13, 14]. In the proposed methods, contribution of each mode is analyzed to obtain a combination of mode estimates to describe the model. Model validation is then carried out using this combination of mode estimates.

The rest of the paper is organized as follows. Section II provides background theory required to implement the proposed methods. Section III discusses methodology of the proposed methods, section IV provides results for different cases validating the effectiveness of the proposed methods, and section V concludes the paper.

2. Background Theory

This section briefly discusses the modeling of power systems under ambient conditions, and the non–paraemtric estimates of auto–covariance function (ACVF) and spectral content of the PMU measurements. The LS–ARMA method, which is used to illustrate the proposed methods, is also briefly described here, the details of which can be found in [15].

*This research was funded by DOE under contracts DE-AC02-05CH11231, DE-SC0012671, and by the University of Wyoming Engineering Initiative.

URI: https://hdl.handle.net/10125/59800
ISBN: 978-0-9981331-2-6

2.1. ARMA model representation of power systems

Under ambient noise conditions, power systems can be represented by linear time-invariant (LTI) ARMA models as shown in Figure 1, where the output $\{y[k]\}_{k=1}^{N}$ represent PMU measurements and input $\{e[k]\}_{k=1}^{N}$ represent an equivalent of the random load variations that continuously drive the system [3]. The random load variation is represented by a zero–mean white noise with variance σ^2.

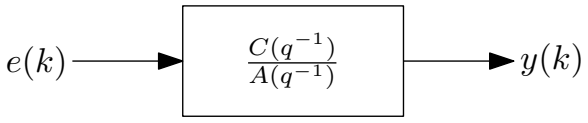


Figure 1. Modeling of power systems under ambient condition

The output measurement, $y[k]$, of the model is given by

$$y[k] = \frac{C(q^{-1})}{A(q^{-1})} e[k], \tag{1}$$

where k is the discrete time index. $C(q^{-1})$ and $A(q^{-1})$ are polynomials defined as
$C(q^{-1}) = c_0 + c_1 q^{-1} + c_2 q^{-2} + ... + c_{n_c} q^{-n_c}$ and
$A(q^{-1}) = 1 + a_1 q^{-1} + a_2 q^{-2} + ... + a_{n_a} q^{-n_a}$,
where q^{-1} is the unit delay operator such that $q^{-m} y[k] = y[k-m]$. The estimates of the electromechanical modes of a system are given by

$$\hat{s}_i = \frac{1}{\Delta T} \log \hat{z}_i, \tag{2}$$

where $\{\hat{z}_i\}_{i=1}^{n_a}$ are the roots of the estimated AR polynomial, $\hat{A}(q^{-1})$, and ΔT is the sampling time period.

2.2. Least–squares ARMA (LS–ARMA)

While any parametric estimation method can be used that can provide estimates of an ARMA model, here a LS–ARMA method is utilized to illustrate the proposed validation methods. LS–ARMA consists of two stages to estimate the ARMA model parameters [15]. In the first stage, an ARMA model is approximated by an AR model of higher order, n_l, where estimates are obtained by solving a linear least squares problem given by

$$\mathbf{Y}\boldsymbol{\alpha} + \mathbf{y} = \mathbf{0}, \tag{3}$$

where

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n_l} \end{bmatrix}^T,$$

$$\mathbf{y} = \begin{bmatrix} y[n_l] & y[n_l+1] & \cdots & y[N] \end{bmatrix}^T \text{ and}$$

$$\mathbf{Y} = \begin{bmatrix} y[n_l-1] & y[n_l-2] & \cdots & y(1) \\ y[n_l-2] & y[n_l-3] & \cdots & y(2) \\ \vdots & \cdots & \cdots & \vdots \\ y[N-1] & y[N-2] & \cdots & y(N-n_l) \end{bmatrix}.$$

The least squares solution of (3) is given by

$$\hat{\boldsymbol{\alpha}} = -\mathbf{Y}^{\dagger}\mathbf{y}, \tag{4}$$

where † denoted pseudo–inverse of a matrix. Then, the noise sequences $\{e[k]\}_{k=n_l+1}^{N}$ are estimated using

$$\hat{e}[k] = y[k] + \sum_{i=1}^{n_l} \hat{\alpha}_i y[k-i]. \tag{5}$$

In the second stage, the coefficients of the AR and MA polynomials are estimated by solving

$$\mathbf{y} + \mathbf{Y}\boldsymbol{\theta} = \mathbf{e}, \tag{6}$$

where

$$\mathbf{Y} = \left[\begin{array}{ccc|ccc} y[L] & \cdots & y(L-n_a+1) & -\hat{e}[L] & \cdots & -\hat{e}(L-n_c+1) \\ y[L+1] & \cdots & y(L-n_a+2) & -\hat{e}[L+1] & \cdots & -\hat{e}(L-n_c+2) \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ y[N-1] & \cdots & y(N-n_a) & -\hat{e}[N-1] & \cdots & -\hat{e}(N-n_c) \end{array}\right],$$

$$\mathbf{y} = \begin{bmatrix} y[L+1] & y[L+2] & \cdots & y[N] \end{bmatrix}^T,$$

$$\mathbf{e} = \begin{bmatrix} e[L+1] & e[L+2] & \cdots & e[N] \end{bmatrix}^T$$

$$\boldsymbol{\theta} = \begin{bmatrix} a_1 & \cdots & a_{n_a} \mid c_1 & \cdots & c_{n_c} \end{bmatrix}^T,$$

and $L > max(n_a, n_c)$. The least squares solution of (6) is given by

$$\hat{\boldsymbol{\theta}} = -\mathbf{Y}^{\dagger}\mathbf{y}, \tag{7}$$

where $\hat{\boldsymbol{\theta}}$ is the parametric estimate of the parameter vector $\boldsymbol{\theta}$. The estimate of the input noise variance is given by

$$\hat{\sigma}^2 = \frac{1}{N-L} (\mathbf{y} + \mathbf{Y}\hat{\boldsymbol{\theta}})^T (\mathbf{y} + \mathbf{Y}\hat{\boldsymbol{\theta}}), \tag{8}$$

where 'T' indicates transpose operation.

2.3. Non–parametric estimates

A non-parametric spectral estimate is obtained by using Welch's periodogram, as described in [15], and given by

$$\hat{\phi}_{NP}(\omega) = \frac{1}{S} \sum_{l=1}^{S} \hat{\phi}_l(\omega), \tag{9}$$

where ω is the frequency in radians/sample,

$$\hat{\phi}_l(\omega) = \frac{1}{MP}\left|\sum_{k=1}^{M} y_l[k]v[k]e^{-i\omega k}\right|^2,$$

$y_l[k] = y((l-1)M + k)$, for $k = 1, 2, ..., M$ and $l = 1, 2, ...S$, denotes the l^{th} data segments each of length M. K is the number of overlapping samples between the adjacent data segments and is usually chosen to be $M/2$. $\mathbf{v}$ is the window vector of length M, and P is the power of the window vector $\mathbf{v}$ given by

$$P = \frac{1}{M}\sum_{k=1}^{M} |v[k]|^2.$$

One should note that other non–parametric spectral estimators could be used such as Blackman–Tukey method [15].

The non–parametric estimate of the ACVF sequences for lag l is given by [15]

$$\hat{R}_{NP}[l] = \frac{1}{N(l)}\sum_{k=l+1}^{N} y[k]y[k-l], \qquad (10)$$

where $N(l) = N$ for the biased ACVF estimate and $N(l) = N - l$ for the unbiased estimate. Because of lower mean–squared error, the biased ACVF estimates will be used.

For the proposed validation method to be useful, it is important that the non–parametric spectral estimate has adequate frequency resolution and small variance. Thus, as is common with mode estimation algorithms, 10 to 20 minutes of data is used and the window size is chosen to be at least 1 minute.

3. Proposed Methods

The proposed methods follow approach similar to that used for validating the mode estimates obtained using Prony analysis in which the contribution of each mode is analyzed to obtain a combination of mode estimates for which the reconstructed signal best matches with the original signal [16]. Using a similar approach, in the first proposed method, parametric and non–parametric estimates of the spectral content of the PMU measurements are compared to visually validate the representation of a system by an estimated ARMA model. In the second proposed method, parametric and non–parametric estimates of the ACVF of the PMU measurements are compared. While the non–parametric estimates are direct functions of the observed data, the parametric estimates are calculated using estimated parametric system model. Therefore, comparing these two estimates can give a fairly good idea about how well these estimated system models agree with the observed data, thereby helping validate estimates of system modes [17]. Before getting into the methodology of the proposed methods, some of the important steps required for implementing the proposed methods are explained in greater detail.

3.1. Calculation of the contribution of each mode to the parametric estimates

In the proposed methods, the contribution of each mode estimate to the spectral content and the ACVF of the PMU measurements is analyzed individually, for which a component corresponding to each of the mode estimates is derived. For this, the parametric spectral density estimate from the PMU measurements in the z–domain, given by

$$\hat{\phi}_P(z) = \hat{\sigma}^2\frac{\hat{C}(z)\hat{C}^*(1/z^*)}{\hat{A}(z)\hat{A}^*(1/z^*)} = \frac{\sum_{k=-n_c}^{n_c}\hat{\gamma}_k z^{-k}}{\sum_{k=-n_a}^{n_a}\hat{\lambda}_k z^{-k}}, \qquad (11)$$

is decomposed into simple partial–fractions [18] given by

$$\hat{\phi}_P(z) = \sum_{i=1}^{n_a}\left(\frac{\hat{r}(i)}{1-\hat{p}(i)z^{-1}} + \frac{\hat{r}_c(i)}{1-\frac{1}{\hat{p}(i)}z^{-1}}\right) \qquad (12)$$

for $n_a > n_c$, where the pole estimates $\{\hat{p}(i)\}_{i=1}^{n_a}$ may be real or complex valued and lie inside the unit–circle in the z–plane and those given by $\{1/\hat{p}(i)\}_{i=1}^{n_a}$ are the ones reflected outside the unit–circle. The term $\hat{r}(i)$ denotes the residue estimate of the i^{th} pole, $\hat{p}(i)$, and $\hat{r}_c(i)$ is the residue estimate of that reflected outside the unit circle and equal to $-\hat{r}(i)$ [18].

The parametric ACVF estimate is obtained by taking inverse z–transform of (12) [15] and given by

$$\begin{aligned}\hat{R}_P[l] &= Z^{-1}\left(\hat{\phi}_P(z)\right)\\ &= \sum_{i=1}^{n_a} Z^{-1}\left(\frac{\hat{r}(i)}{1-\hat{p}(i)z^{-1}} + \frac{\hat{r}_c(i)}{1-(\hat{p}(i))^{-1}z^{-1}}\right)\\ &= \sum_{i=1}^{n_a}\left(\hat{r}(i)(\hat{p}(i))^l u[l] - \hat{r}_c(i)(\hat{p}(i))^{-l}u[-l-1]\right),\end{aligned}$$

where the inverse z–transform of the terms corresponding to the pole estimates $\{\hat{p}(i)\}_{i=1}^{n_a}$ give the ACVF estimates for the positive lags as its region

of convergence (ROC) is outside the circle of radius $|\hat{p}(i)|$ and the pole estimates reflected outside the unit circle contribute to the ACVF estimates for the negative lags as their ROC is inside the circle of radius $|(\hat{p}(i))^{-1}|$ [18]. The ACVF components corresponding to the complex–conjugate pole estimates are combined as follows,

$$\begin{aligned}\hat{r}(i)(\hat{p}(i))^l + \hat{r}^*(i)(\hat{p}^*(i))^l \\ &= \hat{r}(i)\left(\hat{d}_i e^{j\hat{\omega}_i}\right)^l + \hat{r}^*(i)\left(\hat{d}_i e^{-j\hat{\omega}_i}\right)^l \\ &= 2\hat{d}_i^l \times real(\hat{r}(i)e^{jl\hat{\omega}_i}) \\ &= 2|\hat{r}(i)|\hat{d}_i^l \cos(l\hat{\omega}_i + \hat{\beta}_i),\end{aligned}$$

where $\hat{d}_i = |\hat{p}(i)| < 1$, $j = \sqrt{-1}$, $\hat{\omega}_i = \angle\hat{p}(i)$ radians/sample and $\hat{\beta}_i = \angle\hat{r}(i)$ radians. This gives the contribution of each complex mode estimate to the ACVF estimate, which is represented by a damped sinusoid, and is given by

$$\hat{R}_i[l] = \begin{cases} 2|\hat{r}(i)|\hat{d}_i^l \cos(l\hat{\omega}_i + \hat{\beta}_i) & \text{for } l \geq 0 \\ 2|\hat{r}(i)|\hat{d}_i^{-l} \cos(-l\hat{\omega}_i + \hat{\beta}_i) & \text{for } l < 0 \end{cases} \tag{13}$$

Similarly, the contribution of each i^{th} real mode estimate to the parametric ACVF estimate is given by

$$\hat{R}_i[l] = \begin{cases} \hat{r}(i)(\hat{p}(i))^l & \text{for } l \geq 0 \\ \hat{r}(i)(\hat{p}(i))^{-l} & \text{for } l < 0. \end{cases} \tag{14}$$

The contribution of each mode to the parametric estimate of the spectral density is derived by taking discrete–time Fourier transform (DTFT) of the ACVF sequences [15]. Using (14), the contribution of a real mode to the parametric estimate of the spectral density was derived, and is given by

$$\begin{aligned}\hat{\phi}_i(\omega) &= \sum_{l=-\infty}^{\infty} \hat{R}_i[l]e^{-j\omega l} \\ &= \sum_{l=0}^{\infty}\left(\hat{r}(i)(\hat{p}(i))^l e^{-j\omega l}\right) + \sum_{l=-\infty}^{-1}\left(\hat{r}(i)(\hat{p}(i))^{-l} e^{-j\omega l}\right) \\ &= \frac{\hat{r}(i)\left(1-(\hat{p}(i))^2\right)}{1-2\hat{p}(i)\cos(\omega)+(\hat{p}(i))^2}.\end{aligned} \tag{15}$$

Similarly, using (13), the contribution of a complex mode to the parametric estimate of the spectral density was derived, and is given by

$$\begin{aligned}\hat{\phi}_i(\omega) = &\sum_{l=0}^{\infty}\left(2|\hat{r}(i)|\hat{d}_i^l \cos(l\hat{\omega}_i + \hat{\beta}_i)e^{-j\omega l}\right) + \\ &\sum_{l=-\infty}^{-1}\left(2|\hat{r}(i)|\hat{d}_i^{-l} \cos(-l\hat{\omega}_i + \hat{\beta}_i)e^{-j\omega l}\right) \\ = &\frac{|\hat{r}(i)|\cos(\hat{\beta}_i)(1-\hat{d}_i^2) + 2|\hat{r}(i)|\hat{d}_i \sin(\omega-\hat{\omega}_i)\sin(\hat{\beta}_i)}{1-2\hat{d}_i\cos(\omega-\hat{\omega}_i)+\hat{d}_i^2} + \\ &\frac{|\hat{r}(i)|\cos(\hat{\beta}_i)(1-\hat{d}_i^2) - 2|\hat{r}(i)|\hat{d}_i \sin(\omega+\hat{\omega}_i)\sin(\hat{\beta}_i)}{1-2\hat{d}_i\cos(\omega+\hat{\omega}_i)+\hat{d}_i^2},\end{aligned} \tag{16}$$

where the first part of the final form of the equation gives the spectral estimate of a complex mode estimate centered at its positive frequency, given by $\hat{\omega}_i$, and the second part gives the spectral estimate of the mode estimate centered at its negative frequency, given by $-\hat{\omega}_i$.

Thus, a total of $(n_i + n_r)$ components of the ACVF estimates and spectral estimates are obtained, one for each mode estimate, such that $(2n_i + n_r) = n_a$, where n_i is the number of complex pole estimates and n_r is the number of real pole estimates inside the unit circle.

3.2. Sorting of modes

Power systems are of very high order but system dynamics are usually dominated by a modest number of system modes. The model order is therefore chosen large enough to capture the system dynamics, and large enough to not introduce significant bias. The challenge then becomes of how to rank these mode estimates and thus identify the modes that dominate the system dynamics. One possible metric for ranking of the mode estimates would be $\hat{R}_i[0]$ for all i's since this corresponds to the variance of the i^{th} component. Another choice is pseudo–energy of each of the mode estimates, which is basically their contribution to the ACVF estimates as described in [2]. Here, pseudo–energy is defined and then used for ranking because it captures some information about the initial component variance and also damping. This pseudo–energy of each mode estimate is given by the square root of the sum of the squared value of the contribution of the mode estimate to the ACVF estimates,

$$\hat{E}_i = \sqrt{\sum_{l=0}^{\infty} \hat{R}_i[l]\hat{R}_i^*[l]}. \tag{17}$$

Using (13) and (14), pseudo–energy of each i^{th} mode estimate is given by

$$\hat{E}_i = \hat{r}(i)\sqrt{\frac{1}{1-\hat{p}^2(i)}} \quad \text{for real mode}$$

$$= |\hat{r}(i)|\sqrt{2\left(\frac{1}{1-\hat{d}_i^2} + \frac{\cos(2\hat{\beta}_i) - \hat{d}_i^2\cos(2\hat{\omega}_i - 2\hat{\beta}_i)}{1+\hat{d}_i^4 - 2\hat{d}_i^2\cos(2\hat{\omega}_i)}\right)}$$

for complex mode. (18)

3.3. Methodology of the proposed methods

Both proposed methods are multi–step process in which one estimated mode is added in each step, in the decreasing order of their "pseudo energy", to visually analyze the contribution of the mode to the parametric estimates. If the fit between the parametric and non–parametric estimates improves by the addition of a mode estimate then that mode estimate is kept, otherwise discarded. After all the modes have been analyzed, a final combination of the mode estimates is obtained to represent the model. Now, if there is some issue with the data quality or selection of the input variables required by the mode–estimation algorithms, then the parametric estimates, corresponding to the final combination of estimated modes, and the non–parametric estimates will not match that well. For such cases, the model validation methods can be repeated with different data channels or input variables until the match between the parametric and non–parametric estimates is satisfactory. If the match does not improve at all, which can happen for situations such as when the damping ratio of all the modes are high, then a more detailed analysis of the PMU measurements might be required. To summarize, the proposed methods can be carried out using the following steps:

1. Calculate the non–parametric estimates of the ACVF for $l = 0$ to L, where L gives the range of lags, and the signal spectrum for frequency range of interest using (9) and (10);
2. Obtain estimates of the ARMA model parameters, and calculate system modes using (2);
3. Calculate the individual contribution of each of the real and complex modes to the parametric estimates of the spectral content and ACVF using (13)–(16);
4. Calculate pseudo–energy of the modes using (18) and sort them in the decreasing order of their energy;
5. Start with $n_s = 1$, where n_s is the number of selected modes having highest pseudo–energies, and select the mode having the highest pseudo–energy;
6. Add the contribution of the added mode to obtain the parametric estimates of the spectral content and the ACVF corresponding to the selected n_s modes (not including the discarded modes, if any);
7. Match the parametric estimates with the non–parametric estimates and discard the added mode if the contribution of the mode to the parametric estimates does not improve the fit with the non–parametric estimates;
8. In the next and the following steps, increase n_s by 1 and include the mode next in the order of the pseudo–energy;
9. Repeat steps (6)–(8) until all estimated modes are considered to obtain a combination of estimated modes to represent the model;
10. After the final combination of system modes is obtained, the model validation is carried out by comparing the parametric estimates, corresponding to the final combination of mode estimates, with the the non–parametric estimates;
11. If the fit is not satisfactory, then repeat steps (1)–(10) using either different data channels or different input variables.

Some difference between the parametric and the non-parametric estimates is expected as both are merely estimates. As long as the difference is not large, the comparison of the two estimates can give a fairly good idea about the validity of the system representation by the estimated ARMA model.

4. Results and Discussions

The proposed methods were implemented on both simulated data and real world data. Simulated data were generated using the minniWECC model, the details of which can be found in [19]. Some of the dominant modes of the model are NS–A, NS–B and E-W South 1 having frequencies 0.218, 0.372 and 0.510 Hz and damping ratio (ζ) of 7.47%, 4.67% and 8.7093% respectively. Also, real–world data was collected from the WECC system in 2010. A data block of 20 minutes was used to implement the methods.

4.1. Results using simulated data

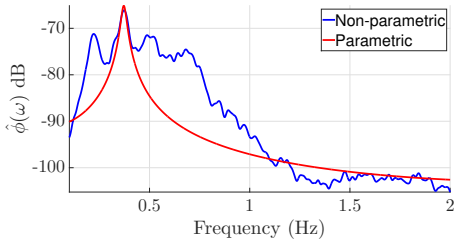
Frequency measurements were used for implementing the proposed methods which were obtained by taking time derivative of the difference of the voltage angle at two different buses. Buses were selected such that the mode of interest had good observability. Data were generated at the rate of 30 samples/sec and were later decimated to 5 samples/sec using approach given in [20] as the frequency of interest usually ranges up to 2 Hz for modal analysis. A number of Monte–Carlo trials were carried out for verifying the effectiveness of the proposed methods. However, as the methods are based on the visual validation, the results for only a single trial are included. Results for most trials were consistent with the set of results included in this paper. For calculating non–parametric spectral estimates, a one minute hamming window was selected.

For the first set of results, data channels were selected such that NS–A, NS–B and ES–South 1 modes had good observability. For case–1, input variables $n_a = 28$, $n_c = 12$ and $n_l = 70$ were chosen to obtain a good match between the parametric and non–parametric estimates. Table 1 shows some of the mode estimates and their normalized pseudo–energy (NPE) listed in the descending order of NPE. Figure 2(a)-2(d) show the results obtained using the first proposed method where the spectral estimates of the signal is plotted in dB scale, and Figure 3(a)-3(d) corresponds to the second method based on ACVF. As explained earlier, the contribution of each mode was analyzed by adding one mode at a time as shown in Figure 2(a) and Figure 3(a) for $n_s = 1$, and Figure 2(b) and Figure 3(b) for $n_s = 2$. Figure 2(d) and Figure 3(d) takes into consideration all the estimated modes for comparison with the results obtained for the final combination shown in Figure 2(c) and Figure 3(c).

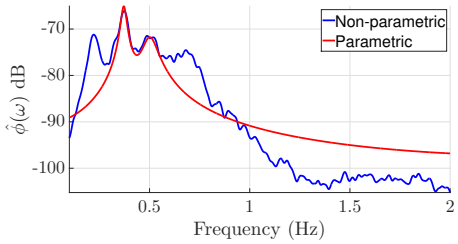
Table 1. Mode-estimates for case–1 and case–2 ranked by normalized pseudo–energy.

Case–1			Case–2		
Freq (Hz)	ζ (%)	N. P. Energy	Freq (Hz)	ζ (%)	N. P. Energy
0.372	3.629	1	0.369	7.523	1
0.501	7.470	0.3607	0.253	18.071	0.6591
0.217	9.440	0.2945	0.578	22.883	0.3754
0.692	7.939	0.1982	0.701	6.980	0.2125
0.792	8.303	0.0149	1.079	8.818	0.0016
0.931	11.960	0.0088	1.493	9.415	8.42E-04
1.540	9.578	5.73E-04	2.112	10.326	3.68E-04
1.854	5.598	4.84E-04	1.867	3.034	3.09E-04
1.218	3.944	2.78E-04			
1.750	3.280	1.04E-04			

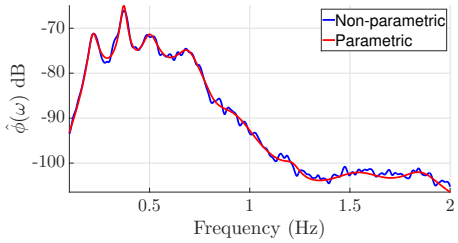
As expected, adding selected modes helped improve the fit between the parametric and non–parametric estimates while the other modes had negligibly small



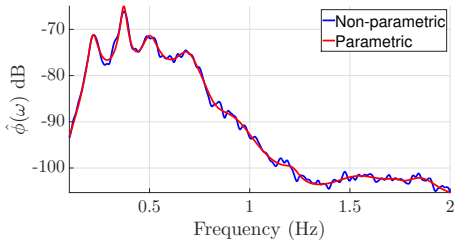
(a) $n_s = 1$



(b) $n_s = 2$



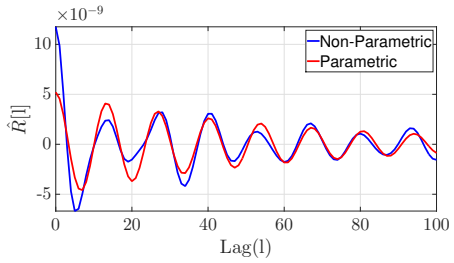
(c) $n_s = 10$



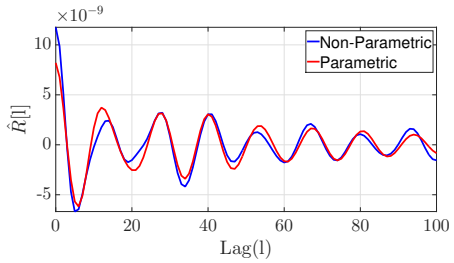
(d) $n_s = 14$

Figure 2. Model validation results for case–1 illustrating the methodology of the proposed method using spectral estimates.

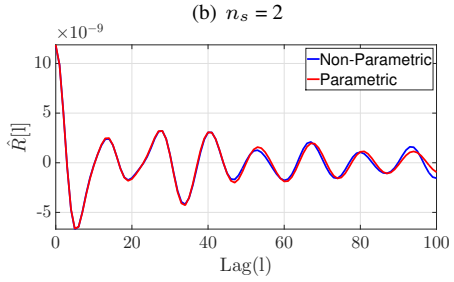
contribution, based on which the final combination of modes were obtained. This combination consisted of the first four modes, shown in Table 1, for the ACVF based method and first ten modes for the other method. The contribution of other mode estimates were negligible as can be seen by comparing Figure 3(c) with Figure 3(d) and Figure 2(c) with Figure 2(d). The final combination of modes based on spectral method consisted of more modes as compared to that obtained by ACVF method, this is because analyzing spectral content, more specifically in log scale, helps evaluate even a much smaller contribution to the parametric



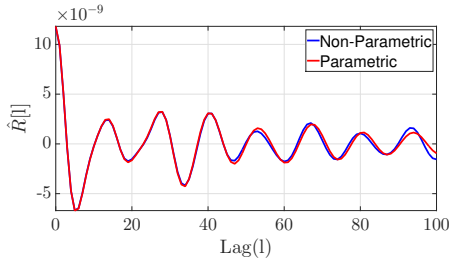
(a) $n_s = 1$



(b) $n_s = 2$



(c) $n_s = 4$



(d) $n_s = 14$

Figure 3. Model validation results for case–1 illustrating the methodology of the proposed method using ACVF estimates.

estimates as can be seen in results. Figure 4 compares the parametric and non–parametric spectral estimates using the first four dominant modes having significant pseudo–energy, and as seen choosing these modes resulted in a good match up to a frequency range of 1 Hz. However, if a good match is needed upto a frequency range of 2 Hz, then additional modes need to be considered that contribute to the spectral densities at higher frequency range. The fit between the parametric and non–parametric estimates for the final combination of the estimated modes were satisfactory, thereby helping validate the estimated system model.

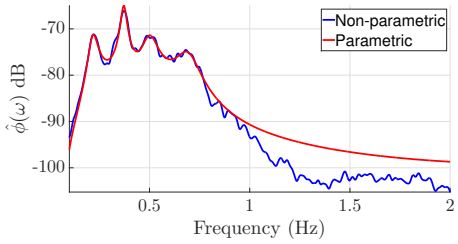
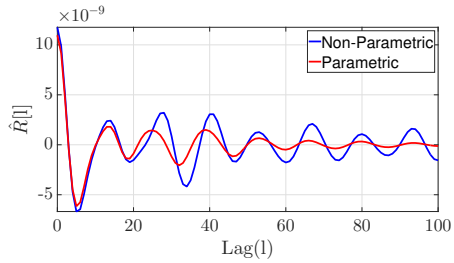
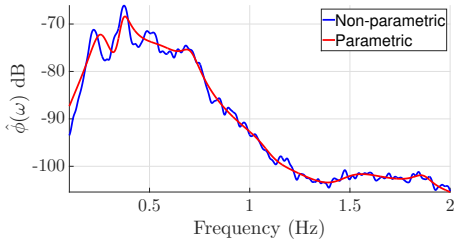


Figure 4. Comparison of parametric and non–parametric spectral estimates for case–1 using four modes having highest energies.



(a) $n_s = 3$



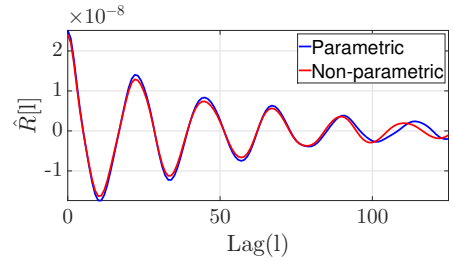
(b) $n_s = 8$

Figure 5. Results for case–2 illustrating the effect of choosing too small model order on the match between the parametric and non–parametric estimates.

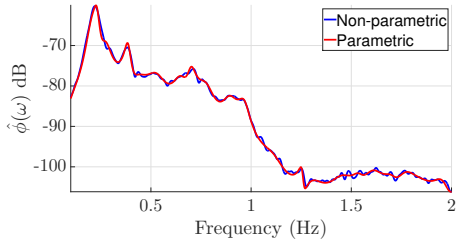
This result can also be verified comparing the estimates of NS–A, NS–B and E–W South 1 modes with their true values.

For case–2, the input variables chosen were $n_a = 16$, $n_c = 14$ and $n_l = 40$ to analyze the performance of the proposed methods when the model order selected was small. The parametric and non–parametric estimates shown in Figure 5(a) and Figure 5(b) do not have a good fit. The mode estimates obtained are shown in Table 1 and the results correspond to the final combination of the mode estimates which consisted of the first three modes for ACVF methods and all eight modes for the method using spectral estimates. In addition to giving biased mode estimates, selection of a reduced model order can also combine modes having close frequencies and hence give inaccurate information about the system modes.

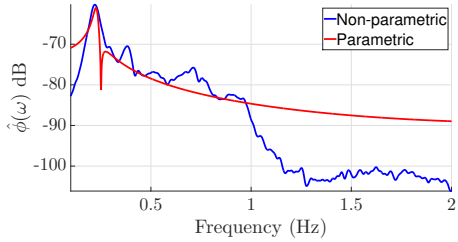
The third case analyzes the performance of the



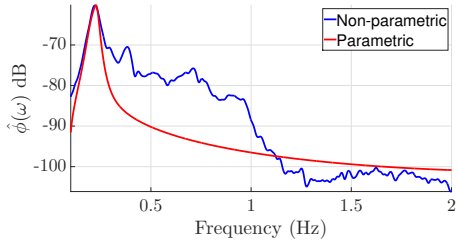
(a) Using the final combination of mode estimates



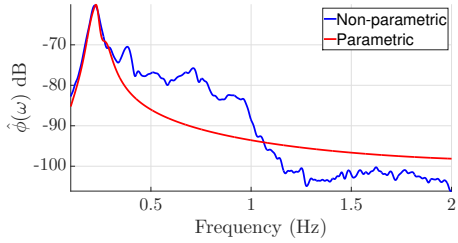
(b) Using the final combination of mode estimates



(c) $n_s = 1$



(d) $n_s = 2$



(e) $n_s = 3$

Figure 6. Results for case–3 illustrating mode–splitting due to over–fitting the data from an excessively high model order.

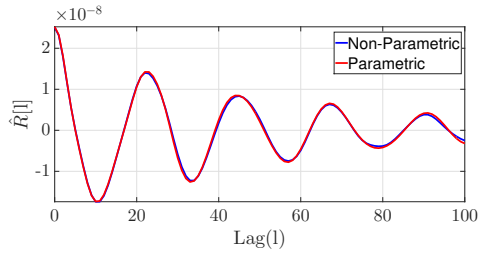
Table 2. Mode-estimates for case–3 and case–4 ranked by normalized pseudo–energy.

Case–3			Case–4		
Freq (Hz)	ζ (%)	N. P. Energy	Freq (Hz)	ζ (%)	N. P. Energy
0.230	6.008	1	0.218	7.613	1
0.202	13.578	0.4351	0.376	4.278	0.0839
0.274	9.585	0.1580	0.255	43.699	0.0450
0.386	3.822	0.1289	0.710	7.871	0.0395
0.701	2.970	0.0327	0.509	10.492	0.0241
0.526	9.067	0.0327	0.942	4.844	0.0088
0.431	10.165	0.0258	0.882	6.897	0.0021
0.648	5.339	0.0230	1.652	8.785	0.0002
0.776	5.221	0.0167	1.988	5.557	5.88E-05
0.894	4.237	0.0074	1.753	2.826	5.54E-05
0.094	52.202	0.0069	1.460	6.761	4.53E-05
0.966	2.942	0.0060	1.1817	2.281	2.86E-05

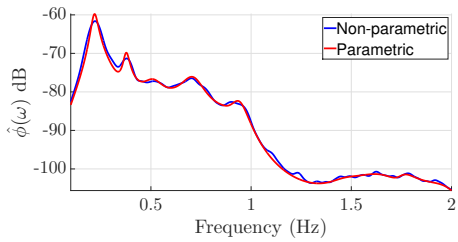
proposed method when the model is over–fitted by choosing a much higher model order. For this case, channels with NS–A mode having high observability were selected and the input variables chosen were $n_a = 56$, $n_c = 24$ and $n_l = 100$. From the results obtained using the final combination of the mode estimates shown in Figure 6(a) and Figure 6(b), it can be seen that the parametric and non–parametric estimates had a good fit. However, analyzing the contribution of the first four modes by adding one mode at a time, as shown in Figure 6(c)-6(e), the mode–splitting of the NS–A mode was observed which is the result of over–fitting of the data from an excessively high model order. This result can also be verified by comparing the NS–A estimates, shown in Table 2, with its true value. For this data set, the input variables $n_a = 28$, $n_c = 20$ and $n_l = 60$ resulted in a good fit between the parametric and non–parametric estimates without over–fitting the model. Some of the mode estimates for this fourth case are given in Table 2 and the results are shown in Figure 7(a) and Figure 7(b) which corresponds to the final combination of the mode estimates. This final combination consisted of the first six modes for the ACVF method and eleven modes for the method based on spectral estimates.

4.2. Results obtained using real PMU measurements

Real–world data were collected from the WECC system in 2010. Due to the confidentiality requirement, limited information is included about the real–world data. The data were pre–processed to remove any outliers, interpolate the missing data and decimate the data to 5 samples/sec. Figure 8(a) and 8(b) show the results obtained for case–1 in which the input variables selected were $n_a = 10$, $n_c = 6$ and $n_l = 35$. These results correspond to the final



(a) $n_s = 6$



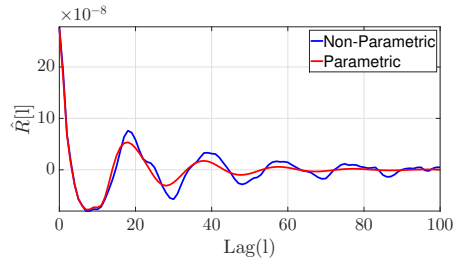
(b) $n_s = 11$

Figure 7. Results for case–4 with parameters selected that gives a good match between the parametric and non–parametric estimates after analyzing the contribution of each mode estimate.

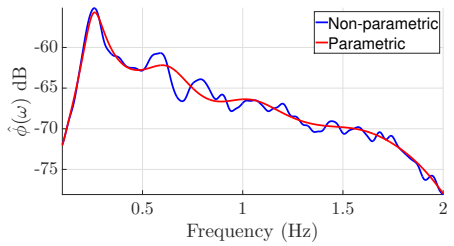
combination of the estimated modes which included all the mode estimates. As seen from the results, the parametric and non–parametric estimates did not have a good fit. Hence, the model validation method was repeated using a different set of input variables, $n_a = 32$, $n_c = 12$ and $n_l = 55$. With this new set of input variables, the parametric and non–parametric estimates matched closely as can be seen in Figure 8(c) and 8(d), thereby validating the representation of the system by the new estimated ARMA model. For case–2, the final combination of the mode estimates consisted of the first nine modes using ACVF method and twelve modes shown in Table 3, for the method based on spectral estimates. Figure 9 compares the parametric and non–parametric spectral estimates using nine dominant modes having large pseudo–energy. For this combination, the parametric and non–parametric estimates had a good match for frequency upto 1.5 Hz, however it did not give a good match for frequency range from 1.5 to 2 Hz as the discarded modes contributed in this frequency range.

5. Conclusion

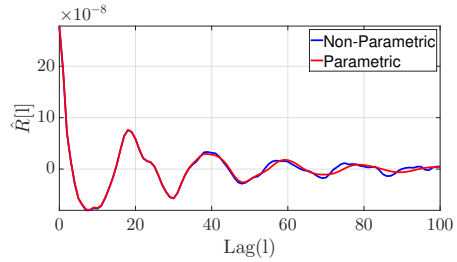
The results obtained using simulated and real–world data illustrated the effectiveness of the proposed methods to validate the mode estimates, obtained under ambient conditions, by comparing parametric and non–parametric estimates of the ACVF and spectral density obtained using parametric estimates of system



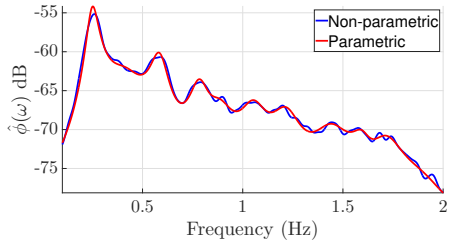
(a) $n_s = 5$



(b) $n_s = 5$



(c) $n_s = 9$



(d) $n_s = 12$

Figure 8. Results obtained using real–world data illustrating the effectiveness of the proposed methods for practical applications.

model and observed data respectively. Based on how well the parametric estimates match with the non–parametric estimates, these methods can help select data channel and input variables required by mode–estimation algorithms. Besides, the proposed methods can also help identify dominant modes of the system based on the analysis of the contribution of each estimated mode to the parametric estimates. Future work will investigate defining metrics to quantify goodness of fit between the parametric and non-parametric estimates. Care must be taken in

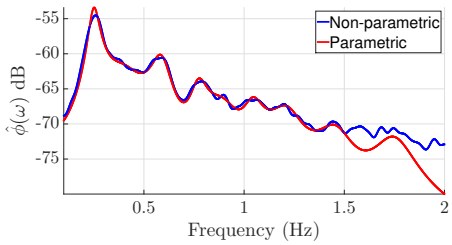


Figure 9. Comparison of parametric and non–parametric spectral estimates using first nine dominant modes.

Table 3. Mode estimates for the real–world data

Case-1			Case-2		
Freq (Hz)	ζ (%)	N. P. Energy	Freq (Hz)	ζ (%)	N. P. Energy
0.254	17.369	1.000	0.249	10.743	1
0.630	20.968	0.341	0.585	7.285	0.2950
1.035	18.985	0.154	0.419	22.547	0.1917
1.599	18.408	0.077	0.784	5.912	0.1247
1.927	15.157	0.022	1.053	6.249	0.0734
			1.216	5.873	0.0652
			0.885	7.635	0.0599
			1.443	6.278	0.0428
			1.715	5.001	0.0327
			1.598	3.205	0.0177
			0	100	0.0100
			1.907	7.419	0.0078

defining such metrics since both are estimates, and the validation comes from consistency between those two estimates since they are formed from very different underlying assumptions. Future work will also focus on validation of the system model using estimated parameters in the presence of forced oscillations and illustrate the method using multiple channels.

Acknowledgement

The authors would like to thank BPA for providing real–world data, and Dr. Daniel Trudnowski from Montana Tech for providing code for the minniWECC model.

References

[1] L. L. Grigsby, *The Electric Power Engineering Handbook*. CRC Press, 2001.

[2] D. J. Trudnowski, J. W. Pierre, N. Zhou, J. F. Hauer, and M. Parashar, "Performance of three mode-meter block-processing algorithms for automated dynamic stability assessment," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 680–690, 2008.

[3] R. W. Wies, J. W. Pierre, and D. J. Trudnowski, "Use of ARMA block processing for estimating stationary low-frequency electromechanical modes of power systems," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 167–173, 2003.

[4] D. Trudnowski and J. Pierre, "Signal processing methods for estimating small-signal dynamic properties from measured responses," in *Inter-Area Oscillations in Power Systems*, pp. 1–36, Springer, 2009.

[5] L. Dosiek, J. W. Pierre, and J. Follum, "A recursive maximum likelihood estimator for the online estimation of electromechanical modes with error bounds," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 441–451, 2013.

[6] G. Liu and V. Venkatasubramanian, "Oscillation monitoring from ambient PMU measurements by frequency domain decomposition," in *2008 IEEE International Symposium on Circuits and Systems*, pp. 2821–2824, May 2008.

[7] S. N. Sarmadi and V. Venkatasubramanian, "Electromechanical mode estimation using recursive adaptive stochastic subspace identification," *IEEE Transactions on Power Systems*, vol. 29, no. 1, pp. 349–358, 2014.

[8] H. Ghasemi, C. Canizares, and A. Moshref, "Oscillatory stability limit prediction using stochastic subspace identification," *IEEE Transactions on Power Systems*, vol. 21, pp. 736–745, May 2006.

[9] Z. Cao and J. W. Pierre, "Electromechanical mode estimation validation using recursive residual whiteness testing," in *North American Power Symposium (NAPS)*, pp. 1–6, IEEE, 2013.

[10] G. K. Pai and J. W. Pierre, "A real-time scheme for validation of an auto-regressive time series model for power system ambient inter-area mode estimation," in *47th Hawaii International Conference on System Sciences (HICSS)*, pp. 2436–2443, IEEE, May 2014.

[11] M. G. Anderson, N. Zhou, J. W. Pierre, and R. W. Wies, "Bootstrap-based confidence interval estimates for electromechanical modes from multiple output analysis of measured ambient data," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 943–950, 2005.

[12] H. Ghasemi and C. A. Canizares, "Confidence intervals estimation in the identification of electromechanical modes from ambient noise," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 641–648, 2008.

[13] A. K. Tangirala, *Principles of System Identification: Theory and Practice*. Crc Press, 2014.

[14] N. Zhou, J. W. Pierre, and J. F. Hauer, "Initial results in power system identification from injected probing signals using a subspace method," *IEEE Transactions on Power Systems*, vol. 21, pp. 1296–1302, Aug 2006.

[15] P. Stoica, R. L. Moses, *et al.*, *Spectral Analysis of Signals*, vol. 452. Pearson Prentice Hall Upper Saddle River, NJ, 2005.

[16] J. F. Hauer, C. J. Demeure, and L. L. Scharf, "Initial results in Prony analysis of power system response signals," *IEEE Transactions on Power Systems*, vol. 5, pp. 80–89, Feb 1990.

[17] L. Ljung, *System identification: Theory for the User*. Prentice-hall, 1987.

[18] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications*. Pearson Prentice Hall, 2007.

[19] D. Trudnowski and J. Undrill, *The MinniWECC System Model*, 2008.

[20] J. W. Pierre, D. J. Trudnowski, and M. K. Donnelly, "Initial results in electromechanical mode identification from ambient data," *IEEE Transactions on Power Systems*, vol. 12, no. 3, pp. 1245–1251, 1997.